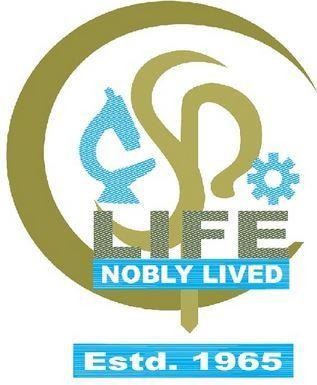
**ST PAUL’S COLLEGE KALAMASSERY**



# MAIN PROJECT REPORT

FRAUD APP DETECTION

*Submitted in partial fulfillment of the requirement for the award of the degree of*

# BSc. Computer Science

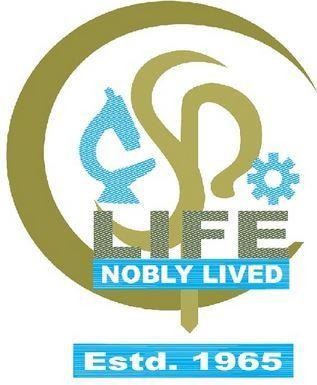
**By Mahatma Gandhi University, Kottayam during 2019-2022**

***Submitted by***

**CN RAKESH NARAYANAN**

**Reg No:190021030197**

**Under the guidance of Mrs. Lincy N.L**

**ST PAUL’S COLLEGE KALAMASSERY**

### **CERTIFICATE**

This is to certify that the main project report entitled **“FRAUD APP DETECTION”** submitted by the candidate **CN Rakesh Narayanan**(RegNo:190021030197) of Batch 2019-2022 during the period of her **Sixth** semester towards the partial fulfillment of requirements for the Bachelor of Science Degree in Computer Science, Mahatma Gandhi University during the academic year 2021-2022 under my supervision and guidance.

### **Mrs. Lincy N. L** **Mrs. Lincy N.L**

**Head of the Department Project Guide**

**Dept. of Computer Science Dept. of Computer Science**

**Submitted for the University Examination held on ……………………………**

**Internal Examiner External Examiner**

**DECLARATION**

I hereby declare that the project work entitled **“B-NEWS”** is the original work done by me during the academic year 2021-2022 under the supervision of Ms. Lincy N.L, Head of the department, Department of Computer Science, St. Paul’s College, Kalamassery during the academic year 2019-2022. This project report is submitted on partial fulfillment of the requirement for the award of Bachelor of Computer Science under MG University.

Date: 15-07-2021 Krishna K P

Place: Kalamassery

# ACKNOWLEDGMENT

First of all, I thank God Almighty for his grace and mercy that enabled me in the finalization of this project.

My sincere thanks to Mrs. Lincy N.L, Head of the Department of Computer Science, for the valuable suggestions and help rendered during the successful completion of this project. I extend my wholehearted thanks for her concern and guidance from the beginning to the end of the project.

I extend my gratitude to all our teachers of the Department of Computer Science for the encouragement given to me during the project period. I express my heartfelt thanks to all my friends and family members for their corporation and support.

**ABSTRACT**

Building upon recent Deep Neural Network architectures, current approaches lying in the intersection of Computer Vision and Natural Language Processing have achieved unprecedented breakthroughs in tasks like automatic captioning or image retrieval. Most of these learning methods, though, rely on large training sets of images associated with human annotations that specifically describe the visual content. In this paper we propose to go a step further and explore the more complex cases where textual descriptions are loosely related to the images. We focus on the particular domain of news articles in which the textual content often expresses connotative and ambiguous relations that are only suggested but not directly inferred from images. We introduce an adaptive CNN architecture that shares most of the structure for multiple tasks including source detection, article illustration and geolocation of articles. Deep Canonical Correlation Analysis is deployed for article illustration, and a new loss function based on Great Circle Distance is proposed for geolocation. Furthermore, we present Breaking News, a novel dataset with approximately 100K news articles including images, text and captions, and enriched with heterogeneous meta-data (such as GPS coordinates and user comments). We show this dataset to be appropriate to explore all aforementioned problems, for which we provide a baseline performance using various Deep Learning architectures, and different representations of the textual and visual features. We report very promising results and bring to light several limitations of current state-of-the-art in this kind of domain, which we hope will help spur progress in the field.

CONTENTS

1. INTRODUCTION 9-11
   1. [OVERVIEW OF THE SYSTEM 10](#_TOC_250012)
   2. [PROBLEM DEFINITION AND OBJECTIVE 11](#_TOC_250011)
2. SYSTEM ANALYSIS 12-28
   1. INTRODUCTION 13
   2. EXISTING SYSTEM 14
   3. PROPOSED SYSTEM 15
   4. FEASIBILITY STUDY 18
   5. [SYSTEM SPECIFICATION 20](#_TOC_250010)
   6. [DATAFLOW DIAGRAM 25](#_TOC_250009)
3. SYSTEM DESIGN 29-42
   1. [LOGICAL DESIGN 30](#_TOC_250008)
   2. [PHYSICAL DESIGN 31](#_TOC_250007)
   3. [INPUT DESIGN 31](#_TOC_250006)
   4. [OUTPUT DESIGN 33](#_TOC_250005)
   5. [DATABASE DESIGN 33](#_TOC_250004)
   6. [TABLE DESIGN 34](#_TOC_250003)
   7. FORMS DESIGN 42
4. SYSTEM DEVELOPMENT 43-45
   1. [CODING 44](#_TOC_250002)
5. SYSTEM TESTING 46-50
   1. [TESTING PROCESS 47](#_TOC_250001)
   2. [METHODS OF TESTING 49](#_TOC_250000)

6.SYSTEM IMPLEMENTATION 51-53

7.MAINTENANCE 54-55

8.CONCLUSION 56-58

8.1 FUTURE SCOPE OF THE PROJECT 58

## 9.APPENDIX 59-65

9.1 SAMPLE FORMS 60

## 10.BIBLIOGRAPHY 66-67

**CHAPTER 1**

**INTRODUCTION**

**1.1 OVERVIEW OF THE SYSTEM**

Building upon recent Deep Neural Network architectures, current approaches lying in the intersection of Computer Vision and Natural Language Processing have achieved unprecedented breakthroughs in tasks like automatic captioning or image retrieval. Most of these learning methods, though, rely on large training sets of images associated with human annotations that specifically describe the visual content. In this paper we propose to go a step further and explore the more complex cases where textual descriptions are loosely related to the images. We focus on the particular domain of news articles in which the textual content often expresses connotative and ambiguous relations that are only suggested but not directly inferred from images. We introduce an adaptive CNN architecture that shares most of the structure for multiple tasks including source detection, article illustration and geolocation of articles. Deep Canonical Correlation Analysis is deployed for article illustration, and a new loss function based on Great Circle Distance is proposed for geolocation. Furthermore, we present BreakingNews, a novel dataset with approximately 100K news articles including images, text and captions, and enriched with heterogeneous meta-data (such as GPS coordinates and user comments). We show this dataset to be appropriate to explore all aforementioned problems, for which we provide a baseline performance using various Deep Learning architectures, and different representations of the textual and visual features. We report very promising results and bring to light several limitations of current state-of-the-art in this kind of domain, which we hope will help spur progress in the field.

**1.2 PROBLEM DEFINITION AND OBJECTIVE**

Building upon recent Deep Neural Network architectures, current approaches lying in the intersection of Computer Vision and Natural Language Processing have achieved unprecedented breakthroughs in tasks like automatic captioning or image retrieval. Most of these learning methods, though, rely on large training sets of images associated with human annotations that specifically describe the visual content. In this paper we propose to go a step further and explore the more complex cases where textual descriptions are loosely related to the images. We focus on the particular domain of news articles in which the textual content often expresses connotative and ambiguous relations that are only suggested but not directly inferred from images. We introduce an adaptive CNN architecture that shares most of the structure for multiple tasks including source detection, article illustration and geolocation of articles. Deep Canonical Correlation Analysis is deployed for article illustration, and a new loss function based on Great Circle Distance is proposed for geolocation. Furthermore, we present BreakingNews, a novel dataset with approximately 100K news articles including images, text and captions, and enriched with heterogeneous meta-data (such as GPS coordinates and user comments). We show this dataset to be appropriate to explore all aforementioned problems, for which we provide a baseline performance using various Deep Learning architectures, and different representations of the textual and visual features. We report very promising results and bring to light several limitations of current state-of-the-art in this kind of domain, which we hope will help spur progress in the field.

**CHAPTER 2**

**SYSTEM ANALYSIS**

**2.1 INTRODUCTION**

Building upon recent Deep Neural Network architectures, current approaches lying in the intersection of Computer Vision and Natural Language Processing have achieved unprecedented breakthroughs in tasks like automatic captioning or image retrieval. Most of these learning methods, though, rely on large training sets of images associated with human annotations that specifically describe the visual content. In this paper we propose to go a step further and explore the more complex cases where textual descriptions are loosely related to the images. We focus on the particular domain of news articles in which the textual content often expresses connotative and ambiguous relations that are only suggested but not directly inferred from images. We introduce an adaptive CNN architecture that shares most of the structure for multiple tasks including source detection, article illustration and geolocation of articles. Deep Canonical Correlation Analysis is deployed for article illustration, and a new loss function based on Great Circle Distance is proposed for geolocation. Furthermore, we present BreakingNews, a novel dataset with approximately 100K news articles including images, text and captions, and enriched with heterogeneous meta-data (such as GPS coordinates and user comments). We show this dataset to be appropriate to explore all aforementioned problems, for which we provide a baseline performance using various Deep Learning architectures, and different representations of the textual and visual features. We report very promising results and bring to light several limitations of current state-of-the-art in this kind of domain, which we hope will help spur progress in the field.

**2.2 EXISTING SYSTEM**

**2.3 PROPOSED SYSTEM**

Building upon recent Deep Neural Network architectures, current approaches lying in the intersection of Computer Vision and Natural Language Processing have achieved unprecedented breakthroughs in tasks like automatic captioning or image retrieval. Most of these learning methods, though, rely on large training sets of images associated with human annotations that specifically describe the visual content. In this paper we propose to go a step further and explore the more complex cases where textual descriptions are loosely related to the images. We focus on the particular domain of news articles in which the textual content often expresses connotative and ambiguous relations that are only suggested but not directly inferred from images. We introduce an adaptive CNN architecture that shares most of the structure for multiple tasks including source detection, article illustration and geolocation of articles. Deep Canonical Correlation Analysis is deployed for article illustration, and a new loss function based on Great Circle Distance is proposed for geolocation. Furthermore, we present BreakingNews, a novel dataset with approximately 100K news articles including images, text and captions, and enriched with heterogeneous meta-data (such as GPS coordinates and user comments). We show this dataset to be appropriate to explore all aforementioned problems, for which we provide a baseline performance using various Deep Learning architectures, and different representations of the textual and visual features. We report very promising results and bring to light several limitations of current state-of-the-art in this kind of domain, which we hope will help spur progress in the field.

**2.4 MODULE DESCRIPTION**

There are 3 modules

1) Admin

2) Reporter

3) User

1. **Admin**

* View reporter details
* View user details
* Add and manage keywords
* Add category and labels
* Add dataset
* View news and reports

1. **Reporter**

* Add report
* View report

1. **User**

* View news
* Search news
* Add review
* View other reviews

# 2.5 FEASIBILITY

After doing the project Automatic Question Paper Generator, study and analyzing all the existing or required functionalities of the system, the next task is to do the feasibility study for the project. All projects are feasible - given unlimited resources and infinite time. Feasibility study includes consideration of all the possible ways to provide a solution to the given problem. The proposed solution should satisfy all the user requirements and should be flexible enough so that future changes can be easily done based on the future upcoming requirements.

### **Economic Feasibility**

This is a very important aspect to be considered while developing a project. We decided the technology based on minimum possible cost factor.

* All hardware and software cost have to be borne by the organization.
* Overall, we have estimated that the benefits the organization is going to receive from the proposed system will surely overcome the initial costs and the later on running cost for system.

### **Technical Feasibility**

This included the study of function performance and constraints that may affect the ability to achieve an acceptable system. For this feasibility study, we studied complete functionality to be provided in the system, as described in the System Requirement Specification (SRS), and check if everything was possible using different type of frontend and backend platforms.

### **Operational Feasibility**

No doubt the proposed system is fully GUI based that is very user friendly and all inputs to be taken all self-explanatory even to a layman. Besides, a proper training has been conducted to let know the essence of the system to the users so that they feel comfortable with new system. As far our study is concerned the clients are comfortable and happy as the system has cut down their loads and doing.

**2.6 SYSTEM SPECIFICATION**

Hardware and software requirements for the installation and smooth functioning of this project could be configured based on the requirements needed by the component of the operating environment that works as front- end system here we suggest minimum configuration for the both hardware and software components.

Working off with this software is requirements concrete on system environments. It includes two phases

* + - Hardware requirements
    - Software requirements